This is the peer reviewed (non-final) version of the following article: Holt DB, Gupta V, Meyer D, Abel NB, Andersen SU, Stougaard J & Markmann K (2015) micro RNA 172 (miR172) signals epidermal infection and is expressed in cells primed for bacterial invasion in *Lotus japonicus* roots and nodules. New Phytologist, 208(1), 241-256. doi: 10.1111/nph.13445, which has been published in final form at <http://dx.doi.org/10.1111/nph.13445>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving."

1 **miR172 signals epidermal infection and is expressed in cells primed for bacterial**
2 **invasion in *Lotus japonicus* roots and nodules**

3 Dennis B. Holt, Vikas Gupta, Dörte Meyer, Nikolaj Abel, Stig U. Andersen, Jens
4 Stougaard and Katharina Markmann
5 Centre for Carbohydrate Recognition and Signalling, Department of Molecular
6 Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, 8000 Aarhus C,
7 Denmark

8 Author for correspondence:

9 *Katharina Markmann*

10 *Tel.: +45 8715 5504;*

11 *Email: kama@mbg.au.dk*

12 **Word count**

13 Total (excluding summary, references and legends): 6769

14 Summary: 199

15 Introduction: 1173

16 Materials and Methods: 242

17 Results: 2532

18 Discussion: 2774

19 Acknowledgements: 110

20 **Figures and Tables**

21 No. of figures: 10 (in colour: Fig. 8)

22 No. of tables: 2

23 **Supporting information**

24 No. of supporting information files: 2

25 (Methods S1, Notes S1 and Figures S1-S3; Supporting Tables S1-S8)

26

27 Summary

- 28 • Legumes interact with rhizobial bacteria to form nitrogen-fixing root nodules.
29 Host signalling following mutual recognition ensures a specific response, but
30 is only partially understood. Focusing on the stage of epidermal infection with
31 *Mesorhizobium loti*, we analysed endogenous small RNAs (sRNAs) of the
32 model legume *Lotus japonicus* to investigate their involvement in host
33 response regulation.
- 34 • We used Illumina sequencing to annotate the *L. japonicus* sRNA-ome and
35 isolate infection-responsive sRNAs, followed by candidate-based functional
36 characterization.
- 37 • Sequences from four libraries revealed 219 novel *L. japonicus* micro RNAs
38 (miRNAs) from 114 newly assigned families, and 76 infection-responsive
39 sRNAs. Unlike infection-associated coding genes such as *NIN*, a miR172
40 isoform strongly accumulated dependent on both Nod factor and compatible
41 rhizobia. Genetics of miR172 induction support a split of epidermal and
42 cortical signalling. *MIR172a* promoter activity followed a previously unseen
43 pattern preceding infection thread progression in epidermal and cortical cells.
44 Nodule-associated miR172a expression was infection-independent,
45 representing the second of two genetically separable activity waves.
- 46 • The combined data provide a valuable resource for further study, and identify
47 miR172 as an sRNA marking successful epidermal infection. We show that
48 miR172 acts upstream of several AP2 transcription factors, and suggest a role
49 in fine-tuning AP2 levels during bacterial symbiosis.

50 Keywords

51 epidermal infection; miRNAs in symbiosis; miR172; *Lotus japonicus*; sRNA
52 sequencing; nodulation symbiosis

53

54

55 **Introduction**

56 Legumes have evolved the ability to respond to nitrogen-fixing bacteria termed
57 rhizobia by forming root nodules. Within these symbiotic organs, the bacteria reside
58 intracellularly, making reduced atmospheric nitrogen available to the plant in
59 exchange for carbohydrates and other nutrients. Following specific mutual
60 recognition, most crop legumes as well as the model hosts *Lotus japonicus* (*Lotus*)
61 and *Medicago truncatula* respond by root hair deformation and the formation of curls,
62 which entrap bacterial microcolonies. The latter are starting points of controlled
63 bacterial entry through tube-like infection threads that traverse epidermal root hair
64 cells and later cortical tissue, where they invade the developing nodule (Desbrosses &
65 Stougaard, 2011). In *Lotus*, symbiont recognition requires the two LysM receptor
66 kinases NFR1 (Madsen *et al.*, 2003) and NFR5 (Madsen *et al.*, 2003; Radutoiu *et al.*,
67 2003), which bind to and mediate the perception of bacterial chito-oligosaccharide
68 Nod factor signals (Broghammer *et al.*, 2012), and determine Nod factor encoded
69 host-symbiont specificity (Radutoiu *et al.*, 2007). NFR5 lacks a functional kinase
70 domain, thus requiring concerted activity with other protein kinases for signal
71 transduction (Madsen *et al.*, 2003; Madsen *et al.*, 2011). Consistently, there is
72 evidence for it forming active heterocomplexes with NFR1 (Madsen *et al.*, 2011) as
73 well as with the leucine-rich repeat receptor-like kinase SYMRK (Endre *et al.*, 2002;
74 Stracke *et al.*, 2002; Antolin-Llovera *et al.*, 2014). Downstream of Nod factor
75 perception, the two cation channels CASTOR and POLLUX (Ane *et al.*, 2004;
76 Imaizumi-Anraku *et al.*, 2005) and three nucleoporins (Kanamori *et al.*, 2006; Saito *et al.*,
77 2007; Groth *et al.*, 2010) are involved in signal transduction prior to the onset of
78 nucleoplasmatic Ca^{++} -oscillations (Ca^{++} -spiking), a response detectable within
79 minutes upon application of Nod factor or compatible rhizobia (Radutoiu *et al.*, 2003;
80 Miwa *et al.*, 2006).

81 The Calcium calmodulin kinase CCaMK (Levy *et al.*, 2004; Tirichine *et al.*, 2006a)
82 and its interactor CYCLOPS (Messinese *et al.*, 2007; Yano *et al.*, 2008a) act
83 downstream of Ca^{++} -spiking (Oldroyd & Downie, 2006; Oldroyd, 2013), which
84 CCaMK is assumed to decode (Singh *et al.*, 2014), triggering subsequent
85 transcriptional reprogramming that ultimately leads to infection thread formation and
86 organogenesis (Madsen *et al.*, 2010; Limpens & Bisseling, 2014). Dominant active

87 phosphomimetic versions of both CCaMK (Yano *et al.*, 2008a) and CYCLOPS
88 (Singh *et al.*, 2014) trigger nodule organogenesis in the absence of rhizobia and each
89 other (Yano *et al.*, 2008a; Singh *et al.*, 2014). Spontaneous nodules can similarly be
90 induced by dominant active forms of the histidine kinase LHK1, a cytokinin receptor
91 acting downstream of the CCaMK/CYCLOPS complex (Tirichine *et al.*, 2007). *Lhk1-1*
92 (formerly *hit1-1*) loss-of-function mutants form only few nodules but feature an
93 overabundance of infection threads in the epidermis of *M. loti* infected roots (Murray
94 *et al.*, 2007).

95 A complex network of transcription factors participates in the control of root hair
96 curling, infection thread formation and nodule organogenesis downstream of Nod
97 factor signalling. Apart from *CYCLOPS*, they include the two GRAS type
98 transcriptional regulators Nodule Signalling Pathway1 and 2 (NSP1/2) (Kalo *et al.*,
99 2005; Smit *et al.*, 2005; Heckmann *et al.*, 2006; Murakami *et al.*, 2006), Nodule
100 Inception (NIN) (Schauser *et al.*, 1999), as well as several AP2/ERF family
101 transcription factors (Middleton *et al.*, 2007; Asamizu *et al.*, 2008; Vernie *et al.*,
102 2008). The *NIN* promoter region was shown to be a direct activation target of
103 phosphorylated CYCLOPS (Singh *et al.*, 2014), and was further shown in
104 *M. truncatula* to be targeted by heterocomplexed NSP1/2 for activation (Hirsch *et al.*,
105 2009). In turn, NIN transcriptionally activates at least two Nuclear factor Y (NF-Y)
106 subunit genes, *NF-YA1* and *NF-YB1*, which have been implicated in the induction of
107 cortical cell divisions that form nodule primordia (Soyano *et al.*, 2013).

108 Within the AP2/ERF transcription factor superfamily, ERF Required for Nodulation1
109 (ERN1) is necessary for nodulation in *M. truncatula* (Middleton *et al.*, 2007), where
110 ERN1 and its homologs ERN2 and ERN3 were found to bind the NF-box, a cis-
111 element driving Nod factor induction of the early nodulation gene *ENOD11*
112 (Andriankaja *et al.*, 2007). In the same species, ERF Required for nodule
113 Differentiation (EFD) is involved in nodule and bacteroid differentiation (Vernie *et al.*,
114 2008). In *Lotus*, ERF1 was identified as a positive regulator of nodulation
115 (Asamizu *et al.*, 2008). So far only two members of the two AP2-domain containing
116 AP2 family were shown to regulate nodulation levels in soybean (Yan *et al.*, 2013;
117 Wang *et al.*, 2014a).

118 Many aspects of the plant response to symbiotic stimuli remain unclear, and small
119 RNAs (sRNAs) are interesting candidates for potential regulators that are not easily
120 identified in forward genetic screens. sRNA guided control of transcript expression,
121 translation and degradation has emerged as a major regulatory strategy in both plants
122 and animals. Apart from affecting core developmental and physiological processes
123 (Chen, 2009; Chuck *et al.*, 2009), sRNA directed silencing plays an essential role in
124 plant responses to biotic and abiotic stresses (Dunoyer *et al.*, 2006; Navarro *et al.*,
125 2006; Ruiz-Ferrer & Voinnet, 2009; Matsui *et al.*, 2013). In recent years, deep
126 sequencing approaches gave access to sRNA populations of diverse plant species,
127 including symbiotic legumes such as soybean (Subramanian *et al.*, 2008; Joshi *et al.*,
128 2010; Kulcheski *et al.*, 2011; Li *et al.*, 2012; Molina *et al.*, 2012; Zabala *et al.*, 2012),
129 *M. truncatula* (Szittyá *et al.*, 2008; Jagadeeswaran *et al.*, 2009a; Lelandais-Briere *et al.*,
130 2009; Devers *et al.*, 2011; Wang *et al.*, 2011; Chen *et al.*, 2012) and *Lotus* (De
131 Luis *et al.*, 2012). Differential sRNA abundances between symbiotic and
132 nonsymbiotic tissues revealed in these studies suggest a role for sRNAs in root
133 symbioses.

134 Plant miRNAs cover a size range of circa 19-24nt (Ramachandran & Chen, 2008;
135 Chen, 2009) and are generated from imperfectly complementary hairpin precursors
136 predominantly through DCL1 action (Bologna & Voinnet, 2014). They typically
137 control mRNA target activity through cleavage-mediated decay or translational
138 inhibition (Brodersen *et al.*, 2008; Brodersen & Voinnet, 2009). Although miRNAs
139 constitute a minor fraction of regulatory sRNAs, they have received most attention
140 among sRNAs in plants, including legumes. A few representatives have been
141 functionally associated with rhizobial symbiosis (Combier *et al.*, 2006; Boualem *et al.*,
142 2008; Li *et al.*, 2010; Turner *et al.*, 2013; Yan *et al.*, 2013; Wang *et al.*, 2014a) or
143 shown to regulate transcription factors known to be involved in the latter (Combier *et al.*,
144 2006; Devers *et al.*, 2011; De Luis *et al.*, 2012; Lauressergues *et al.*, 2012). Most
145 studies identify symbiosis-related miRNAs in later symbiotic stages or focus on their
146 function in nodule development. As sRNAs are fast-evolving molecules (Chen, 2009;
147 Cuperus *et al.*, 2011) that can significantly alter plant response patterns to external
148 stimuli (Ruiz-Ferrer & Voinnet, 2009; Matsui *et al.*, 2013), we saw potential for their
149 involvement also in early symbiotic responses. Our study aims to promote the
150 understanding of symbiotic signalling events by characterising response patterns of

small host RNAs during initiation of nitrogen-fixing root symbiosis. We use Illumina deep sequencing to trace sRNA dynamics during the first steps of legume interactions with compatible rhizobia, and further characterise miR172a, a conserved miRNA that shows two waves of induction during symbiosis and signals successful epidermal infection.

Materials and Methods

Plants for sequencing and qPCR assays were *Lotus japonicus* L. ecotype Gifu wild type, *nfr1-1* (Schauser *et al.*, 1998; Radutoiu *et al.*, 2003; Sandal *et al.*, 2006), *nfr5-1* (Madsen *et al.*, 2003; Sandal *et al.*, 2006), *symrk-1* (Schauser *et al.*, 1998; Stracke *et al.*, 2002), *ccamk-2* (Demchenko *et al.*, 2004; Kistner *et al.*, 2005; Tirichine *et al.*, 2006b), *cyclops-2* (Schauser *et al.*, 1998; Sandal *et al.*, 2006; Yano *et al.*, 2008a; Perry *et al.*, 2009), *lhk1-1* (formerly *hit1-1*) (Murray *et al.*, 2007), *nin-1* (Schauser *et al.*, 1999; Sandal *et al.*, 2006), *nsp1-1* and *nsp2-3* (Heckmann *et al.*, 2006), *har1* (Wopereis *et al.*, 2000; Krusell *et al.*, 2002), *sst1* (Krusell *et al.*, 2005) *sen1* (Suganuma *et al.*, 2003) and *snf1* (Tirichine *et al.*, 2006a). Plants were prepared and grown as detailed in the Supporting Methods online and inoculated at indicated time points with *M. loti* R7A wild type, *M. loti* R7A *nodc* (Rodpothong *et al.*, 2009) or *M. loti* MAFF303099 expressing *DsRED* (Maekawa *et al.*, 2009) bacteria. For Nod factor treatment, purified *M. loti* R7A Nod factor dissolved in 50% isoamyl alcohol was applied at a concentration of 10^{-7} Mol/liter. Control plants were mock-treated with equal amounts of the same medium ($\frac{1}{2}$ strength Broughton and Dilworth, 1 mM KNO₃) used for suspending bacteria and diluting Nod factor solutions. Oligonucleotide sequences used for reverse transcription, quantitative PCR and for cloning are listed in Table S1. Please see Methods S1 online for additional experimental details.

Results

Deep sequencing of sRNA populations from *Lotus* roots

Using Illumina deep sequencing we analysed sRNA populations from 10-day-old *Lotus* roots that were either inoculated with *M. loti* or mock-treated (Table 1). Roots were harvested three days post inoculation at the stage of bacterial entry into root epidermal cells. Four independent libraries were sequenced, representing two biological replicates (Table 1) and yielding circa 34 mio reads (Fig. 1). Following

183 adapter filtering and trimming, a global abundance cut-off of 10 was applied, which
 184 removed potentially erroneous sequences, including most of those containing
 185 unidentified nucleotide (N) positions, from the dataset. About 378 thousand unique
 186 sequences of 19-24nt length passed these criteria (Fig. 1, Table S2) and were used for
 187 onward profiling analysis as well as mapping- or prediction-based annotation steps as
 188 outlined in Fig. 1. Size fractionation revealed that all libraries contain a majority of
 189 24nt reads. These were highly diverse and individual reads tended to be lowly
 190 abundant, contributing to a higher proportion of 24nt reads in the unique compared to
 191 redundant dataset (Fig. 2, Table S2). 21 and 22nt fractions represented the second and
 192 third most abundant size classes, respectively. Especially in the 21nt fraction, few
 193 unique sequences reached high accumulation levels, resulting in a higher
 194 representation in the redundant compared to the unique dataset (Fig. 2, Table S2).
 195 About 80% of unique pass filter sequences mapped to one or more positions on the
 196 *L. japonicus* genome v2.5 (Sato *et al.*, 2008). Of these, 6.8% were annotated as
 197 repeats, and another 0.4% mapped to positions on the genome of *M. loti* (Fig. 1).

198 ***Lotus* root sRNAs respond to bacterial infection**

199 To identify rhizobia-induced regulation of sRNAs we generated profiles comparing
 200 the normalized abundance of each unique sequence across libraries. There were
 201 distinct differences between root sRNA populations from infected compared to mock-
 202 treated plants. Using the globally applied cutoff of 10 sequence counts isolated a large
 203 pool of potentially symbiosis responsive sequences. To separate relevant candidates
 204 from potential background, we applied a raised abundance cut-off of 100. This
 205 identified 14 sequences that were at least 100x less abundant in inoculated compared
 206 to mock-treated roots, and 62 sequences that showed a 100x or more increase in
 207 abundance in inoculated libraries of both biological replicates (Table 2). To compare
 208 global expression patterns between treatments and replicates, we plotted unique
 209 sequences to reflect their normalized abundances in the four libraries (Fig. 3). The
 210 plots revealed much higher divergence between treatments (Fig. 3a) than between
 211 biological replicates (Fig. 3b).

212 **Infection responsive sRNAs are enriched in novel and conserved miRNAs**

213 We predicted novel miRNAs present in our dataset using miRDeep-P (Yang & Li,
 214 2011), which was more efficient than miRDeep2 (Friedländer *et al.*, 2012) in

215 annotating conserved miRNAs identified by mapping to the miRBase v19 mature
216 miRNA registry (Kozomara & Griffiths-Jones, 2011) (Fig. S1).

217 A total of 167 unique sequences (Table S3a) mapping to 219 precursor gene loci (Fig.
218 1, Table S3b) predicted as miRNAs by miRDeep-P within the filtered sequence pool
219 showed no homology to known miRNAs (miRBase v19) and passed the criteria put
220 forth by the miRBase miRNA registry (<http://www.mirbase.org/>). They were
221 annotated and registered as isoforms of 114 novel miRNA families. An additional 65
222 miRDeep-P predicted miRNA sequences (28% of a total of 232 predicted, confirmed
223 miRNA sequences) showed homology to previously annotated miRNAs (miRBase
224 v19) (Fig. 1, Table S3c).

225 In comparison, only 8 (out of 154) candidate miRNA sequences predicted by
226 miRDeep2 mapped to miRBase v19 entries. There was little overlap between
227 miRDeep2 and miRDeep-P predictions, and we thus considered only the latter in our
228 further analyses.

229 Of the predicted novel miRNAs, 6 were at least 10-fold more, and 6 at least 10-fold
230 less abundant in libraries from *M. loti*- compared to mock-treated roots of both
231 biological replicates (Table S4).

232 To identify symbiosis-responsive conserved miRNAs that were not co-predicted by
233 the miRDeep-P algorithm, we mapped the filtered dataset to mirBase v19 using a
234 three-mismatch-criterion. 1272 unique sequences mapped to previously annotated
235 plant miRNAs or miRNA*s (1253) or were classified as homologous to known
236 miRNAs by miRBase (19) (Table S5). Of these, 29 were at least 10-fold more
237 abundant in inoculated compared to mock-treated samples, with a further 10 showing
238 an opposite abundance pattern in both biological replicates (Table S4). Both novel
239 and conserved miRNA candidates were thus clearly enriched among sequences
240 responsive to rhizobial inoculation, where they contributed 0.81% and 2.62%,
241 respectively, while making up only 0.06% and 0.34% of the total pool of pass filter
242 sequences. Relatively seen, this enrichment is thus particularly pronounced among
243 novel miRNAs, suggesting that some of these specifically evolved in the context of
244 rhizobial endosymbiosis. Like the majority of symbiosis-regulated sRNAs in our
245 libraries, most responsive miRNA candidates belong to families that also contain
246 nonresponsive isoforms (Table S4).

247 **miR172a levels are strongly enhanced upon infection**

248 Two sequences showed particularly strong upregulation in *M. loti* infected compared
 249 to mock-treated roots (Fig. 3a; black circles). One of these originated from a
 250 multigene family generating several sRNA sequences with divergent abundance
 251 patterns in mock-treated versus infected roots. The expression specificities were thus
 252 difficult to confirm using hybridization-based approaches, as co-detection of
 253 homologs obscured the results. This candidate was therefore not analysed further. The
 254 second (Fig. 3a, arrows; Table S4a) was identified as a member of the widely
 255 conserved *MIR172* family, represented in our dataset with 19 isoforms from at least 6
 256 distinct loci (Table S4c). We named it miR172a.

257 qPCR analysis targeting the mature miR172a sequence confirmed its regulatory
 258 pattern on the same RNA samples used for sequencing library generation, yet at lower
 259 amplitudes (Fig. 4a). This was equally the case for another infection-responsive
 260 sequence, a miR171 isoform (Fig. 4b).

261 To correlate miR172a expression dynamics with the infection process, we then
 262 investigated abundance levels of this miRNA during the first few hours and days upon
 263 *M. loti* infection (Fig. 4). miR172a levels remained low within hours of inoculum
 264 application (Fig. 4c), but were markedly and progressively higher in infected
 265 compared to mock-treated roots at 2, 3 and 4 dpi (Fig. 4d).

266 **miR172a regulation depends on Nod factor signalling, but is independent of the** 267 **cytokinin receptor LHK1**

268 To position miR172a function relative to the processes of infection thread formation
 269 and nodule organogenesis, we tested its inducibility in various mutant backgrounds at
 270 a time point preceding nodule primordium formation. The latter is apparent from
 271 about 4 dpi where miR172a is significantly more abundant than at 3 dpi (Fig. 4d), and
 272 we thus tested miR172a levels at 3 dpi to avoid co-detection of possible
 273 organogenesis-related miR172a regulation. *L. japonicus* plants impaired in the Nod
 274 factor receptors NFR1 and NFR5, *nfr1-1* and *nfr5-1* respectively, failed to display
 275 altered miR172a levels upon *M. loti* inoculation (Fig. 5a). A miR172a response was
 276 similarly undetectable in *symrk-1* and *ccamk-2* loss-of-function mutants, both of
 277 which show aberrant root hair responses but do not initiate infection thread formation
 278 or organogenesis (Stracke *et al.*, 2002; Tirichine *et al.*, 2006a) (Fig. 5a). Finally,

279 *cyclops-2* mutants, lacking infection threads but with unimpaired root hair curling or
280 organogenesis initiation (Yano *et al.*, 2008b) did not show infection-related miR172a
281 induction.

282 We then examined miR172a levels in *lhk1-1* mutants, which develop an
283 overabundance of epidermal infection threads, while nodules are only induced
284 occasionally (Murray *et al.*, 2007). *M. loti* infected roots of this mutant showed
285 similarly enhanced levels of miR172a as wild type plants (Fig. 5b).

286 qPCR analysis of *nin-1* mutant roots revealed a complete loss of miR172a induction
287 (Fig. 5b). As *lhk1-1* loss-of-function mutants were not impaired in miR172a
288 responsiveness (Fig. 5b), this suggests that NIN is active independent of LHK1 at this
289 stage of symbiosis development. To confirm this result, we tested whether *NIN*
290 similarly retained responsiveness in *lhk1-1* plants. Indeed, *NIN* transcripts were
291 induced to higher than wild type levels upon infection after three days (Fig. 5c).

292 Similar to NIN, the two GRAS type transcription factors NSP1 and NSP2 are thought
293 to be involved in both infection and organogenesis control. We investigated their
294 influence on miR172a induction and found a strong attenuation of miR172a
295 abundance increase in loss-of-function *nsp1-1* and *nsp2-3* mutants (Fig. 5b). This
296 result is consistent with NSP1/2 action upstream of miR172, while the observed
297 residual activation suggests the involvement of one or more additional factors at the
298 NSP1/2 hierarchical level. The observation that *NIN* induction is strongly attenuated
299 in the absence of functional NSP1 (Fig. 5c) suggests that NIN acts downstream of the
300 latter towards miR172a induction. Alongside NSP1, we observed a partial
301 dependence of NIN regulation on functional CYCLOPS, whereas a full wild type-like
302 *NIN* response was seen in *nsp2* mutant plants (Fig. 5c).

303 **miR172a induction in roots requires both bacteria and Nod factors**

304 The dependence of infection-related miR172a induction on Nod factor signalling
305 genes suggest a role of Nod factors in triggering this response. We thus treated plants
306 with purified Nod factors isolated from the *M. loti* strain R7A. At three days post
307 treatment, miR172a showed no traceable response (Fig. 6a), but the Nod factor
308 inducible transcription factor genes *NIN* (Schauser *et al.*, 1999) and *NF-YAI*
309 (Høgslund *et al.*, 2009; Soyano *et al.*, 2013) were strongly upregulated (Fig. 6b). This
310 indicates that Nod factor alone is insufficient for miR172a induction. To test whether

other bacterial associated factors are instead sufficient to trigger the response, we then inoculated plants with bacteria carrying a transposon insertion in the NodC gene and unable to produce Nod factors (Rodpohong *et al.*, 2009). Like purified Nod factor, but unlike wild type bacteria of the same strain R7A, nodC mutants did not trigger enhanced miR172a levels (Fig. 6a). Interestingly, we saw strain specific differences in miR172a induction levels using wild type bacteria of either R7A or MAFF 303099 strains (Fig. 5 and 6).

miR172a levels are enhanced in nodules independent of their functionality

To investigate whether miR172a regulation is linked specifically to early, pre-organogenesis stages of infection in *Lotus*, we used qPCR to analyse miR172a abundances during later stages of symbiosis. At 3 wpi, miR172a was several thousand times more abundant in nodules compared to mock-treated root tissue (Fig. 7). To test whether miR172a is correlated with nitrogen fixing activity of bacteroids, we tested its abundance levels in nodules of *sst1* and *sen1* mutants which form fully infected nodules unable to fix nitrogen (Suganuma *et al.*, 2003; Krusell *et al.*, 2005). miR172a abundance in *sen1* and *sst1* nodules is strongly enhanced as compared to uninoculated roots of the same genotypes (Fig. 7), though to a lesser extent than in functional wild type nodules (Fig. 7).

To test if nodule-associated miR172a expression is retained in the absence of rhizobial infection, we monitored miR172a abundance in spontaneous nodules of *snf1*, a mutant line carrying an auto-active form of CCaMK (Tirichine *et al.*, 2006a). miR172a accumulates in spontaneous nodules compared to uninoculated *snf1* roots, though abundance levels are less than in infected *snf1* or wild type nodules (Fig. 7).

MIR172a expression coincides with infection-primed root and nodule cells

To resolve the spatial and temporal expression of *MIR172a*, we transformed roots with a 1.8 kb *MIR172a* promoter region fused to the β -glucuronidase (GUS) reporter gene (*pMIR172a:GUS*). Transgenic control roots as well as mock-treated *pMIR172a:GUS* roots showed no apparent GUS staining. *M. loti*-inoculated *pMIR172a:GUS* roots displayed no visible GUS-staining at 1 dpi (Fig. 8a,b), but at 2 dpi single root hairs were stained, often in combination with confined groups of epidermal and outer cortical cells or a combination of both (Fig. 8c,d). GUS stained root hairs showed curling, indicating responsiveness to *M. loti* (Fig. 8d). At 3 and 4

343 dpi, inoculated roots showed GUS-staining in several well-defined spots along the
 344 roots (Fig. 8e-h). Cross sections revealed strongly stained sites of apparent
 345 meristematic activity in the root cortex, likely nodule primordia (Fig. 8h).

346 At 7 dpi, developing nodules are infected but generally not yet efficient in nitrogen
 347 fixation, whereas roots harvested at 3 wpi display fixing nodules of various
 348 developmental stages (Fig. 8i). 7 dpi nodules and immature nodules harvested at 3
 349 wpi showed GUS staining in and immediately around central, infected tissue (Fig.
 350 8j,k). With increasing maturity the expression intensity lessened in the infected centre
 351 (Fig. 8l), and was confined to the pericentric nodule parenchyme in fully
 352 differentiated nodules (Fig. 8m). To trace the progression of *MIR172a* expression
 353 relative to *M. loti* infection during symbiosis development, we used *pMIR172a* to
 354 express a nuclear localized tYFP marker (*pMIR172a:tYFPnls*). Consistent with the
 355 GUS results, *pMIR172a:tYFPnls* roots inoculated with *dsRED* expressing *M. loti*
 356 displayed confined epidermal and cortical groups of yellow fluorescent nuclei in the
 357 direct vicinity of bacterial infections (Fig. 8n,o). Cross sections of infected roots
 358 revealed tYFP expression in non-dividing cells several cell layers ahead of infection
 359 thread progression (Fig. 8o). Expression patterns seen in immature (Fig. 8p) and
 360 mature (Fig. 8q) nodules were consistent with GUS reporter observations. Fluorescent
 361 nuclei included both roundish and elongated ones, the latter being characteristic of
 362 infected cells, indicating that expression was active in infected as well as uninfected
 363 cells of the central nodule tissue. Mature nodules (Fig. 8q) displayed a ring of weakly
 364 expressing cells surrounding the infected tissue.

365 **miR172a targets AP2-like transcription factors**

366 To investigate the biological role of miR172a-target nodes, we predicted mRNA
 367 targets using the psRNA target algorithm (Dai & Zhao, 2011) to search *L. japonicus*
 368 gene models (Sato *et al.*, 2008). Using an e-value cutoff of 2, five AP2/EREBP
 369 transcription factors were predicted subjects of miR172a directed cleavage (Table
 370 S6). Three of these show high levels of amino acid identity to Arabidopsis RAP2-
 371 7/TOE1 (55%, 58% and 55%, respectively), whereas two are closely related to the
 372 Arabidopsis APETALA 2 (AP2) with 66% and 70% amino acid identity, respectively
 373 (Table S6). The predicted targets are thus referred to here as RAP2-7-like1, RAP2-7-
 374 like2, RAP2-7-like3, AP2-like1 and AP2-like2.

To test these mRNAs for miR172a directed endonucleolytic cleavage, we analysed an in-house degradome deep sequencing dataset prepared from infected root and nodule tissue harvested at 3 and 21 dpi, respectively (Holt, 2014). All five candidates showed accumulation of degradome read 5' ends in the centre of the miR172a complementary sites, suggesting cleavage at a single dominant position (Fig. 9). *Lotus* transcriptome data (Høgslund *et al.*, 2009) indicate that three of the five target mRNAs, namely *RAP2-7-like1*, *AP2-like1* and *AP2-like2*, are significantly less abundant in nodules compared to uninoculated roots (Table S7). This is complementary to what the same dataset shows for the *MIR172a* precursor or what we observed for the mature miR172a (Fig. 7). qPCR analysis confirmed these tendencies for all three targets (Fig. 10a-c). At early stages of infection, between 1-4 dpi, qPCR analysis revealed no significant expression changes for these targets (Fig. 10d-f).

Misexpression of *MIR172a* or its targets suggests a fine-tuning role

To reveal roles of miR172a during infection and nodulation, we changed the abundances of mature miR172a or the target mRNAs in transgenic *Lotus* roots using a set of promoters that either follow *pMIR172a* expression (*pMIR172a*, *pLYS3*), drive expression, in contrast to *pMIR172*, primarily in roots but less so in nodules (*p35S*), or effect overexpression across most cell types (*pUBQ1*). In neither setup, infection threads or nodules on transgenic roots showed significant qualitative or quantitative abnormalities (Table S8; Notes S1 and Supporting Figures S1 and S2 online). These combined observations suggest a role for the miR172a-target nodes in fine-tuning symbiotic establishment.

Discussion

Infection responsive sRNAs are enriched in both novel and conserved miRNAs

We analysed *Lotus* sRNA responses during infection with rhizobial bacteria, focusing on the stage of mutual recognition accompanied by the initiation, but preceding the development of symbiotic root nodules. Although sRNA pools have been isolated from nodules and other tissues of various legumes, little is known about their involvement at this early, predominantly epidermal symbiotic stage. sRNAs showing differential abundances were enriched in both conserved and novel miRNAs, suggesting that miRNA directed fine-tuning of gene regulation is involved in the initiation phase of symbiosis development. A recent study investigating *Lotus* sRNA

407 responses using Roche 454 sequencing (De Luis *et al.*, 2012) did not identify
408 significant changes in miRNA populations after 3 hours or 3 dpi. Adapted library
409 preparation protocols and increased sequencing depth present a possible basis for the
410 improved resolution. In contrast, a targeted approach found several soybean miRNAs
411 including miR159 and miR393 isoforms, to respond to rhizobial recognition at a pre-
412 infection stage (Subramanian *et al.*, 2008). Representatives of these families were
413 among the candidates linked to infection in our dataset. Candidates correlating
414 negatively with symbiont recognition in soybean include miR160 (Subramanian *et al.*,
415 2008), which is represented in our dataset with several infection-downregulated
416 isoforms and has been positively associated with pathogen responses in other plants
417 (Padmanabhan *et al.*, 2009). These examples suggest parallels between sRNA
418 responses during symbiont recognition and those accompanying the following
419 epidermal infection. Members of several miRNA families represented in our
420 infection-responsive pool have been identified in various other legume tissues,
421 including nodules, and the positively responsive miR172, miR156 and miR171 have
422 been shown to accumulate in nodule organs compared to root tissues in *M. truncatula*
423 and/or *Lotus* (Lelandais-Briere *et al.*, 2010; De Luis *et al.*, 2012). These three
424 candidates have also been shown to be functionally involved in rhizobial symbiosis at
425 the stage of nodulation (Yan *et al.*, 2013; Hofferek *et al.*, 2014; Wang *et al.*, 2014a;
426 Wang *et al.*, 2014b), but none have so far been linked to epidermal infection. To date,
427 although certain miRNAs have been functionally associated with legume bacterial or
428 fungal root symbiosis, possible roles of siRNAs in symbiosis are still emerging.
429 Indirect evidence comes from the involvement of miR482 in soybean nodulation, a
430 miRNA that targets defence related resistance gene receptor kinases (Li *et al.*, 2010;
431 Shivaprasad *et al.*, 2012). These have, in turn, been shown to act as sources for
432 phased siRNAs in several plant species including legumes (Lelandais-Briere *et al.*,
433 2009; Zhai *et al.*, 2011; Shivaprasad *et al.*, 2012). Interestingly, two miR482 isoforms
434 are among the infection-upregulated miRNAs in our dataset, suggesting miRNA-
435 triggered abundance control of resistance gene transcripts/proteins as an early
436 necessity during symbiosis establishment.

437 Members of at least two further infection-upregulated miRNA families in our dataset
438 trigger the generation of phased siRNAs from their targets in plants, *MIR159* targeting
439 MYB transcription factors and *MIR393* cleaving TIR/AFB auxin receptor targets (Fei

440 *et al.*, 2013). As siRNAs appear to form a substantial part of the inoculation-
 441 responsive pools observed in our dataset, their functional involvement during early as
 442 well as later stages of legume nodulation represents an exciting basis for further
 443 studies.

444 **Different response patterns within conserved miRNA families suggest their**
 445 **functional diversification**

446 miR172, miR171 and other miRNA families represented in our pool of infection-
 447 responsive sequences (Tables S4) have been shown to be involved in controlling core
 448 developmental processes in plants including the nonsymbiotic *Arabidopsis* (Llave *et al.*,
 449 2002; Chen, 2004; Wu *et al.*, 2009; Yu *et al.*, 2012; Curaba *et al.*, 2013),
 450 suggesting that miRNAs involved in bacterial root symbiosis have evolved from
 451 diverse nonsymbiotic contexts. This possibility is underlined by the observation that
 452 responsiveness to rhizobial inoculation was generally limited to a subset of isoforms
 453 within families (Table S5). In line with these findings, miRNA families were found to
 454 contain isoforms displaying distinct response patterns to conditions such as salt stress
 455 or symbiotic nodulation in *M. truncatula* (Lelandais-Briere *et al.*, 2009).

456 **miR172a depends on Nod factor signalling, but is independent of the cytokinin**
 457 **receptor *LHK1***

458 miR172a clearly stood out as the most strongly infection-regulated sRNA in our
 459 dataset. miR172 isoforms have previously been detected in legume roots or nodules
 460 (Subramanian *et al.*, 2008; Jagadeeswaran *et al.*, 2009b; Lelandais-Briere *et al.*, 2009;
 461 Wang *et al.*, 2009; Valdes-Lopez *et al.*, 2010; De Luis *et al.*, 2012), but
 462 responsiveness to infection has not been documented. In soybean, miR172 was
 463 recently shown to trigger increased nodule numbers when ectopically expressed in
 464 roots (Yan *et al.*, 2013). The steep increase in *Lotus* miR172a abundance at 3 dpi,
 465 where infection thread formation and epidermal invasion have commenced, points to
 466 a role at the stage of epidermal infection. To clarify the developmental and genetic
 467 context of miR172a activity, we made use of a series of mutants impaired in signal
 468 transduction and transcriptional reprogramming (Høgslund *et al.*, 2009; Madsen *et al.*,
 469 2010) that are phenotypically distinct from wild type plants at 3 dpi. Mutants
 470 impaired in Nod factor receptor genes *NFR1* and *NFR5* were nonresponsive for
 471 miR172a, as were all tested downstream signalling genes required for both rhizobial
 472 and AM fungal infection in *Lotus*, including *SYMRK* (Stracke *et al.*, 2002), *CCaMK*

(Tirichine *et al.*, 2006a) and *CYCLOPS* (Yano *et al.*, 2008a). These genes have been associated either directly or indirectly with the induction of both epidermal infection and organogenesis (Madsen *et al.*, 2010). The abortion of miR172a induction in *cyclops*-deficient mutants forming nodule primordia but no infection threads supports miR172a's specific association with infection rather than organogenesis. This scenario is further supported by its retainment in the histidine kinase *LHK1* loss-of-function mutant *lhk1-1*, a cytokinin perception mutant that develops an overabundance of epidermal infection events, but rarely nodules (Murray *et al.*, 2007). Our observation that miR172a induction is independent of *LHK1* but depends on the transcription factor *NIN* at 3 dpi is interesting in light of previous findings where *NIN* expression control depended on *LHK1* at 7 dpi, placing it downstream of the latter at this stage (Murray *et al.*, 2007). This supports a model where *NIN* acts independently of *LHK1* in epidermal infection but downstream of it at organogenesis-linked cortical stages.

The genetic dependencies of miR172a induction combined with those of *NIN* (Fig. 5) open a new perspective on symbiosis signalling (Fig. S3), revealing that (a) full *NIN* induction requires both *CYCLOPS* and *NSP1* activity; that (b) *NSP1* is, in contrast to *Medicago* *NSP1* (Hirsch *et al.*, 2009), involved in *NIN* induction either without or with only facultative involvement of *NSP2*, which in its absence might be replaced by a close homolog showing similar regulation (Høgslund *et al.*, 2009); that (c) *NIN*, although necessary for miR172a induction, is not sufficient to trigger it, since *NIN* induction is unimpaired in *nsp2-3* mutants where miR172a expression is strongly attenuated, and is retained in *cyclops* mutants where miR172a induction is entirely off; and that (d) full miR172a induction requires *CYCLOPS* as well as *NSP2* mediated activities other than *NIN* regulation (Fig. S3). In line with this latter hypothesis, work by Hirsch *et al.* (2009) suggests that *NSP1* and *NSP2* act as a complex in *M. truncatula* to regulate induction of additional symbiosis-inducible genes, namely *Early Nodulation 11* (*ENOD11*) and the ERF transcription factor *ERN1*. It is tempting to speculate on a regulation of miR172a-targeted AP2 transcription factors at this level that could then positively influence abundance levels of their sRNA regulator via a feedback mechanism. We suggest that signalling downstream of *NIN* towards miR172a induction occurs in a threshold dependent

505 manner, since no intermediate miR172a induction is seen in the absence of functional
506 *NIN* or *CYCLOPS* (Fig. S4).

507 The observation that spontaneous nodule organogenesis induction by an auto-active
508 *CYCLOPS* version requires both *NSP1* and *NSP2* (Singh *et al.*, 2014) points to them
509 acting downstream of *CYCLOPS*-mediated *NIN* activation (Singh *et al.*, 2014). This
510 scenario is consistent with miR172a induction strictly depending on the activities of
511 *CYCLOPS* and *NIN*, but does not explain the strong reduction of *NIN* activation in
512 *nsp1* mutants that we observed. In combination with *cyclops* loss-of-function mutants
513 retaining partial *NIN* activity (Singh *et al.*, 2014, Yano *et al.* 2006 and this study), our
514 data support a contribution of both *NSP1*, as described for *Medicago* (Hirsch *et al.*,
515 2009), and *CYCLOPS* to *NIN* activation with miR172a induction depending on these
516 combined activities. In such a scenario, miR172a induction would strictly depend on
517 the activities of *CYCLOPS* and *NIN*, but only partially on *NSP1/2* mediated
518 downstream gene activation including that of *NIN*.

519 In summary, infection-related miR172a activity is downstream of all four tested
520 transcription factors, but upstream of or parallel to *LHK1* mediated cytokinin
521 perception in *Lotus*.

522 A second miRNA upregulated in response to rhizobial infection, miR171c, has been
523 shown to directly regulate *NSP2* involving endonucleolytic cleavage in both *Lotus*
524 (De Luis *et al.*, 2012) and *M. truncatula* (Devers *et al.*, 2011; Laressergues *et al.*,
525 2012). The genetics of miR172a regulation thus reveal the existence of at least two
526 nodes of miRNA activity during infection-related signalling in legumes.

527 **miR172a is a sRNA marker for infection at the pre-organogenesis stage**

528 To analyse which factors are involved in triggering miR172a upregulation in *Lotus*,
529 we made use of purified bacterial Nod factor signals as well as a bacterial mutant
530 unable to produce the latter. The surprising finding was that both Nod factor and
531 bacterial presence are required for early miR172a activity, suggesting its dependency
532 on multiple factors linked to the onset of symbiotic infection. In line with these
533 findings, the intensity of miR172a induction in response to *M. loti* infection depended
534 on the bacterial strain used: A variant of *M. loti* MAFF 303099 notable for its
535 vigorous infection properties (personal observation) consistently triggered higher
536 miR172a levels than the strain R7A. Both produce the same dominant Nod factor,

537 suggesting that either other bacterial factors or molecular events following the
 538 resulting infection contribute qualitatively to miR172a induction. The molecular
 539 cause underlying these observations will be interesting to investigate further.

540 **Independent events control epidermal and nodule-associated waves of miR172a** 541 **activity**

542 In contrast to the pre-organogenesis stage, nodule-associated *MIR172a* expression
 543 occurs in the absence of infection if organogenesis is triggered via an autoactive
 544 version of *CCaMK*, as is the case in *snf1* mutant plants. This is in line with previous
 545 findings (De Luis *et al.*, 2012), and confirms that independent signalling events
 546 control epidermal and nodule-associated miR172a activity: Ca^{++} spiking, decoded by
 547 *CCaMK*, and downstream responses can be achieved by Nod factor application alone
 548 (Felle *et al.*, 1998; Felle *et al.*, 1999), but these events are insufficient to trigger
 549 miR172a activation during the epidermal phase of symbiosis.

550 Consistent with miR172a induction in spontaneous nodules, infected but non-
 551 functional nodules formed by loss-of-function mutants of *SST1* (Krusell *et al.*, 2005)
 552 and *SEN1* (Hakoyama *et al.*, 2012) equally show enhanced miR172a compared to
 553 uninoculated roots. Yet in all cases, the nodule-levels of miR172a activity were
 554 reduced compared to wild type functional nodules. The latter observation suggests a
 555 positive feedback mechanism involved in miR172a regulation that depends on nodule
 556 infection and functionality. This is in line with the multiple regulatory levels,
 557 including feedback loops, participating in miR172 abundance control in the context of
 558 Arabidopsis floral timing and organ patterning (Zhu & Helliwell, 2011).

559 ***MIR172a* expression is associated with infected or infection-primed cells**

560 *pMIR172a* activity seems to overlap with and pre-trace the path of infection
 561 progression, starting at the epidermis and expanding towards the inner cortex prior to
 562 initiation of cell divisions forming nodule primordia. Later in the developmental
 563 process, expression associates with developing nodule primordia (Fig. 8). Among
 564 symbiotic genes this is an unusual expression pattern, and the pre-infection
 565 demarcation of cells may be used as a tool enabling future identification and
 566 characterization of cells with infection competence and/or cells supporting infection
 567 thread formation. Along the same line, *pMIR172:GUS* expression patterns may be
 568 used as a basis for isolating mutants impaired in preparing epidermal, cortical or

569 nodule cells for infection. The abundance increase in spontaneous *snfl* nodules
 570 demonstrates that in nodule organs, while *pMIR172a* expression is limited to the
 571 infected centre and its close vicinity in wild type nodules (Fig. 5), this pattern is
 572 independent of bacterial presence.

573 miR172 has previously been found to be present in nodules (Lelandais-Briere *et al.*,
 574 2009; Wang *et al.*, 2009; Valdes-Lopez *et al.*, 2010; De Luis *et al.*, 2012) and was
 575 detected in the infection zone of *M. truncatula* nodules (Lelandais-Briere *et al.*,
 576 2009), suggesting a similar role in both determinate and indeterminate legume
 577 nodules. We observe that nodule-associated *pMIR172a* activity ceases as nodules
 578 mature. In combination with the root activity pattern at early stages of symbiosis, this
 579 suggests an involvement of miR172a in priming cells for infection and possibly in the
 580 processes of infection and differentiation into sink-tissue rather than in maintenance
 581 and function of bacteroids within active, mature nodules.

582 **miR172a targets *AP2-like* transcription factors**

583 The miR172 complementary site is present in *AP2* homologs of diverse land plant
 584 lineages (Kim *et al.*, 2006; Shigyo *et al.*, 2006), suggesting conserved regulatory
 585 mechanisms. miR172 directed abundance control of AP2 and its homologs is essential
 586 in determining floral organ identity and promoting vegetative phase change and
 587 flowering in Arabidopsis (Aukerman & Sakai, 2003; Chen, 2004; Wu *et al.*, 2009).
 588 Similar roles have been described in monocots such as maize (Lauter *et al.*, 2005) and
 589 rice (Zhu *et al.*, 2009) as well as in *Lotus* (Yamashino *et al.*, 2013), indicating a high
 590 level of functional conservation.

591 We identified five *APETALA 2 (AP2)-like* transcription factors as targets of miR172
 592 isoforms, and three of these showed mRNA abundance levels reciprocal to miR172a
 593 in nodules. The latter observation suggests regulation of target transcripts through
 594 endonucleolytic cleavage at least in nodule tissue. Prior to nodule development,
 595 miR172a induction was not paralleled by a traceable decrease in target mRNAs.
 596 miR172-target nodes have been shown in Arabidopsis to involve translational
 597 inhibition as primary control mode (Aukerman & Sakai, 2003; Chen, 2004), and this
 598 may similarly prevail over mRNA degradation at the stage of epidermal infection in
 599 *Lotus*. An alternative explanation is signal dilution due to the relatively small fraction

600 of root cells responding to *M. loti* infection and displaying miR172a peaks upon
601 inoculation.

602 Functional implications of miR172 isoforms are looked at as complementary to those
603 of miR156 regulating SPL (Squamosa Promoter Binding Protein Like) transcription
604 factors (Wu *et al.*, 2009). Both miRNA families affect diverse developmental
605 phenotypes associated with phase change (Wu *et al.*, 2009). Interestingly, in addition
606 to miR172 and along with miR159 isoforms, which have equally been implicated in
607 flowering time control (Achard *et al.*, 2004), several miR156 isoforms are among the
608 top regulated conserved miRNAs in our dataset (Table S4).

609 While these combined observations link bacterial root symbiosis to core
610 developmental processes, the roles of these miRNA families and their targets in
611 symbiosis are still unclear and will be interesting to investigate further. Altering
612 miR172a or target abundances in *Lotus* did not reveal apparent phenotypical effects,
613 perhaps reflecting limitations of the chimeric plant assay used: while roots with
614 diverse transgene insertion sites can be tested, variation is high, and subtle phenotypes
615 are difficult to detect. Together with previous findings where miR172 abundance has
616 been linked to quantitative, but no apparent qualitative features of symbiosis in
617 soybean (Yan *et al.*, 2013; Wang *et al.*, 2014a), the elusive role in *Lotus* suggests that
618 the identified miR172a-target nodes hold subtle and possibly redundant roles in fine-
619 tuning symbiosis-related plant responses. As for plant development as a whole, this
620 may be a general feature of the bulk miRNAs involved in symbiosis.

621 **Conclusions**

622 Our results suggest that regulatory sRNAs, and miRNAs in particular, play an
 623 important part in the early steps of legume infection with rhizobial bacteria. Several
 624 infection responsive miRNAs, including miR172a, belong to families implicated in
 625 conserved developmental processes, indicating that symbiosis evolution involved
 626 functional diversification of miRNA directed gene regulation. miR172a is strongly
 627 regulated during the early stages of symbiosis dependent on infection, and expression
 628 outlines infection thread progression. In the signalling cascade leading to infection
 629 thread formation, miR172a acts downstream of *NIN*, which is required but not
 630 sufficient for its response, in line with miR172a induction requiring CYCLOPS and
 631 NSP1 activities other than *NIN* regulation. In contrast to *NIN*, miR172a is
 632 nonresponsive to Nod factor signals alone, making it a unique marker for successful
 633 symbiosis initiation. At later stages of symbiosis, *MIR172a* expression is limited to
 634 central nodule cells, and this second wave of regulation is independent of infection
 635 and nodule functionality. We show that miR172a regulates at least three AP2
 636 transcription factors involved in key developmental processes in Arabidopsis. While
 637 the roles of these regulatory nodes in symbiosis are still unclear, they are likely subtle,
 638 helping to fine-tune plant responses to bacterial infection downstream of recognition
 639 and Nod factor signalling.

640 **Acknowledgements**

641 We are grateful to Simon Moxon and Frank Schwach (University of East Anglia,
 642 Norwich, UK) for helpful discussion and support concerning the UEA sRNA tools
 643 platform, and John Sullivan (University of Otago, New Zealand) and Nicolaj N.
 644 Maolanon (University of Copenhagen, Denmark) who kindly supplied Nod factor
 645 preparations. We thank Linards Klavins for excellent technical support and Yasuyuki
 646 Kawaharada for helpful discussion. The sequencing data is available at Gene
 647 Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>), record GSE35885.
 648 miRNAs predicted by miRDeep-P (Yang & Li, 2011) have been deposited in the
 649 miRBase miRNA registry (<http://www.mirbase.org/>). This work was supported by the
 650 Danish National Research Foundation grant no. DNRF79 and the ERC Advanced
 651 Grant 268523.

652

653

654 **References**

- 655 **Achard P, Herr A, Baulcombe DC, Harberd NP. 2004.** Modulation of floral
656 development by a gibberellin-regulated microRNA. *Development* **131**: 3357-3365.
- 657 **Andrianakaja A, Boisson-Dernier A, Frances L, Sauviac L, Jauneau A, Barker**
658 **DG, de Carvalho-Niebel F. 2007.** AP2-ERF transcription factors mediate Nod factor
659 dependent *MtENOD11* activation in root hairs via a novel cis-regulatory motif. *Plant*
660 *Cell* **19**: 2866-2885.
- 661 **Ane JM, Kiss GB, Riely BK, Penmetza RV, Oldroyd GE, Ayax C, Levy J,**
662 **Debelle F, Baek JM, Kalo P, Rosenberg C, Roe BA, Long SR, Denarie J, Cook**
663 **DR. 2004.** *Medicago truncatula DMII* required for bacterial and fungal symbioses in
664 legumes. *Science* **303**: 1364-1367.
- 665 **Antolin-Llovera M, Ried MK, Parniske M. 2014.** Cleavage of the SYMBIOSIS
666 RECEPTOR-LIKE KINASE ectodomain promotes complex formation with Nod
667 factor receptor 5. *Curr Biol* **24**: 422-427.
- 668 **Asamizu E, Shimoda Y, Kouchi H, Tabata S, Sato S. 2008.** A positive regulatory
669 role for *LjERF1* in the nodulation process is revealed by systematic analysis of
670 nodule-associated transcription factors of *Lotus japonicus*. *Plant Physiol* **147**: 2030-
671 2040.
- 672 **Aukerman MJ, Sakai H. 2003.** Regulation of flowering time and floral organ
673 identity by a microRNA and its *APETALA2*-like target genes. *Plant Cell* **15**: 2730-
674 2741.
- 675 **Bologna NG, Voinnet O. 2014.** The diversity, biogenesis, and activities of
676 endogenous silencing small RNAs in Arabidopsis. *Annu Rev Plant Biol* **65**: 473-503.
- 677 **Boualem A, Laporte P, Jovanovic M, Laffont C, Plet J, Combier JP, Niebel A,**
678 **Crespi M, Frugier F. 2008.** MicroRNA166 controls root and nodule development in
679 *Medicago truncatula*. *Plant J* **54**: 876-887.
- 680 **Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P,**
681 **Yamamoto YY, Sieburth L, Voinnet O. 2008.** Widespread translational inhibition
682 by plant miRNAs and siRNAs. *Science* **320**: 1185-1190.
- 683 **Brodersen P, Voinnet O. 2009.** Revisiting the principles of microRNA target
684 recognition and mode of action. *Nat Rev Mol Cell Biol* **10**: 141-148.
- 685 **Broghammer A, Krusell L, Blaise M, Sauer J, Sullivan JT, Maolanon N, Vinther**
686 **M, Lorentzen A, Madsen EB, Jensen KJ, Roepstorff P, Thirup S, Ronson CW,**
687 **Thygesen MB, Stougaard J. 2012.** Legume receptors perceive the rhizobial
688 lipochitin oligosaccharide signal molecules by direct binding. *Proc Natl Acad Sci U S*
689 *A* **109**: 13859-13864.

- 690 **Chen L, Wang T, Zhao M, Tian Q, Zhang WH. 2012.** Identification of aluminum-
691 responsive microRNAs in *Medicago truncatula* by genome-wide high-throughput
692 sequencing. *Planta* **235**: 375-386.
- 693 **Chen X. 2004.** A microRNA as a translational repressor of APETALA2 in
694 Arabidopsis flower development. *Science* **303**: 2022-2025.
- 695 **Chen X. 2009.** Small RNAs and their roles in plant development. *Annu Rev Cell Dev*
696 *Biol* **25**: 21-44.
- 697 **Chuck G, Candela H, Hake S. 2009.** Big impacts by small RNAs in plant
698 development. *Curr Opin Plant Biol* **12**: 81-86.
- 699 **Combier JP, Frugier F, de Billy F, Boualem A, El-Yahyaoui F, Moreau S, Vernie**
700 **T, Ott T, Gamas P, Crespi M, Niebel A. 2006.** MtHAP2-1 is a key transcriptional
701 regulator of symbiotic nodule development regulated by microRNA169 in *Medicago*
702 *truncatula*. *Genes Dev* **20**: 3084-3088.
- 703 **Cuperus JT, Fahlgren N, Carrington JC. 2011.** Evolution and functional
704 diversification of *MIRNA* genes. *Plant Cell* **23**: 431-442.
- 705 **Curaba J, Talbot M, Li Z, Helliwell C. 2013.** Over-expression of microRNA171
706 affects phase transitions and floral meristem determinancy in barley. *BMC Plant Biol*
707 **13**: 6.
- 708 **Dai X, Zhao PX. 2011.** psRNATarget: a plant small RNA target analysis server.
709 *Nucleic Acids Res* **39**: W155-159.
- 710 **De Luis A, Markmann K, Cognat V, Holt DB, Charpentier M, Parniske M,**
711 **Stougaard J, Voinnet O. 2012.** Two microRNAs linked to nodule infection and
712 nitrogen-fixing ability in the legume *Lotus japonicus*. *Plant Physiol* **160**: 2137-2154.
- 713 **Demchenko K, Winzer T, Stougaard J, Parniske M, Pawlowski K. 2004.** Distinct
714 roles of *Lotus japonicus* *SYMRK* and *SYM15* in root colonization and arbuscule
715 formation. *New Phytol* **163**: 381-392.
- 716 **Desbrosses GJ, Stougaard J. 2011.** Root nodulation: a paradigm for how plant-
717 microbe symbiosis influences host developmental pathways. *Cell Host & Microbe* **10**:
718 348-358.
- 719 **Devers EA, Branscheid A, May P, Krajinski F. 2011.** Stars and symbiosis:
720 microRNA- and microRNA*-mediated transcript cleavage involved in arbuscular
721 mycorrhizal symbiosis. *Plant Physiol* **156**: 1990-2010.

- 722 **Dunoyer P, Himber C, Voinnet O. 2006.** Induction, suppression and requirement of
 723 RNA silencing pathways in virulent *Agrobacterium tumefaciens* infections. *Nat Genet*
 724 **38:** 258-263.
- 725 **Endre G, Kereszt A, Kevei Z, Mihacea S, Kalo P, Kiss GB. 2002.** A receptor
 726 kinase gene regulating symbiotic nodule development. *Nature* **417:** 962-966.
- 727 **Fei Q, Xia R, Meyers BC. 2013.** Phased, secondary, small interfering RNAs in
 728 posttranscriptional regulatory networks. *Plant Cell* **25:** 2400-2415.
- 729 **Felle HH, Kondorosi E, Kondorosi A, Schultze M. 1998.** The role of ion fluxes in
 730 Nod factor signalling in *Medicago sativa*. *Plant J* **13:** 455-463.
- 731 **Felle HH, Kondorosi E, Kondorosi A, Schultze M. 1999.** Nod factors modulate the
 732 concentration of cytosolic free calcium differently in growing and non-growing root
 733 hairs of *Medicago sativa* L. *Planta* **209:** 207-212.
- 734 **Friedländer MR, Mackowiak SD, Li N, Chen W, Rajewsky N. 2012.** miRDeep2
 735 accurately identifies known and hundreds of novel microRNA genes in seven animal
 736 clades. *Nucleic Acids Res* **40:** 37-52.
- 737 **Groth M, Takeda N, Perry J, Uchida H, Draxl S, Brachmann A, Sato S, Tabata**
 738 **S, Kawaguchi M, Wang TL, Parniske M. 2010.** *NENA*, a *Lotus japonicus* homolog
 739 of *Sec13*, is required for rhizodermal infection by arbuscular mycorrhiza fungi and
 740 rhizobia but dispensable for cortical endosymbiotic development. *Plant Cell* **22:**
 741 2509-2526.
- 742 **Gupta V, Markmann K, Pedersen CN, Stougaard J, Andersen SU. 2012.**
 743 shortran: a pipeline for small RNA-seq data analysis. *Bioinformatics* **28:** 2698-2700.
- 744 **Hakoyama T, Niimi K, Yamamoto T, Isobe S, Sato S, Nakamura Y, Tabata S,**
 745 **Kumagai H, Umehara Y, Brossuleit K, Petersen TR, Sandal N, Stougaard J,**
 746 **Udvardi MK, Tamaoki M, Kawaguchi M, Kouchi H, Suganuma N. 2012.** The
 747 Integral Membrane Protein SEN1 is Required for Symbiotic Nitrogen Fixation in
 748 *Lotus japonicus* Nodules. *Plant Cell Physiol* **53:** 225-236.
- 749 **Heckmann AB, Lombardo F, Miwa H, Perry JA, Bunnewell S, Parniske M,**
 750 **Wang TL, Downie JA. 2006.** *Lotus japonicus* nodulation requires two GRAS
 751 domain regulators, one of which is functionally conserved in a non-legume. *Plant*
 752 *Physiol* **142:** 1739-1750.
- 753 **Hirsch S, Kim J, Munoz A, Heckmann AB, Downie JA, Oldroyd GE. 2009.**
 754 GRAS proteins form a DNA binding complex to induce gene expression during
 755 nodulation signaling in *Medicago truncatula*. *Plant Cell* **21:** 545-557.

- 756 Hofferek V, Mendrinna A, Gaudé N, Krajinski F, Devers E. 2014. miR171h
757 restricts root symbioses and shows like its target *NSP2* a complex transcriptional
758 regulation in *Medicago truncatula*. *BMC Plant Biol* **14**.
- 759 Høgslund N, Radutoiu S, Krusell L, Voroshilova V, Hannah MA, Goffard N,
760 Sanchez DH, Lippold F, Ott T, Sato S, Tabata S, Liboriussen P, Lohmann GV,
761 Schauser L, Weiller GF, Udvardi MK, Stougaard J. 2009. Dissection of symbiosis
762 and organ development by integrated transcriptome analysis of *Lotus japonicus*
763 mutant and wild-type plants. *PLoS One* **4**: e6556.
- 764 Holt DB (2014) *Small RNAs involved in bacterial root symbiosis and its regulation in*
765 *Lotus japonicus*. PhD Thesis, Aarhus University, Aarhus, Denmark.
- 766 Imaizumi-Anraku H, Takeda N, Charpentier M, Perry J, Miwa H, Umehara Y,
767 Kouchi H, Murakami Y, Mulder L, Vickers K, Pike J, Downie JA, Wang T, Sato
768 S, Asamizu E, Tabata S, Yoshikawa M, Murooka Y, Wu GJ, Kawaguchi M,
769 Kawasaki S, Parniske M, Hayashi M. 2005. Plastid proteins crucial for symbiotic
770 fungal and bacterial entry into plant roots. *Nature* **433**: 527-531.
- 771 Jagadeeswaran G, Saini A, Sunkar R. 2009a. Biotic and abiotic stress down-
772 regulate miR398 expression in Arabidopsis. *Planta* **229**: 1009-1014.
- 773 Jagadeeswaran G, Zheng Y, Li YF, Shukla LI, Matts J, Hoyt P, Macmil SL,
774 Wiley GB, Roe BA, Zhang W, Sunkar R. 2009b. Cloning and characterization of
775 small RNAs from *Medicago truncatula* reveals four novel legume-specific microRNA
776 families. *New Phytol* **184**: 85-98.
- 777 Joshi T, Yan Z, Libault M, Jeong DH, Park S, Green PJ, Sherrier DJ, Farmer A,
778 May G, Meyers BC, Xu D, Stacey G. 2010. Prediction of novel miRNAs and
779 associated target genes in *Glycine max*. *BMC Bioinformatics* **11 Suppl 1**: S14.
- 780 Kalo P, Gleason C, Edwards A, Marsh J, Mitra RM, Hirsch S, Jakab J, Sims S,
781 Long SR, Rogers J, Kiss GB, Downie JA, Oldroyd GE. 2005. Nodulation signaling
782 in legumes requires *NSP2*, a member of the GRAS family of transcriptional
783 regulators. *Science* **308**: 1786-1789.
- 784 Kanamori N, Madsen LH, Radutoiu S, Frantescu M, Quistgaard EM, Miwa H,
785 Downie JA, James EK, Felle HH, Haaning LL, Jensen TH, Sato S, Nakamura Y,
786 Tabata S, Sandal N, Stougaard J. 2006. A nucleoporin is required for induction of
787 Ca²⁺ spiking in legume nodule development and essential for rhizobial and fungal
788 symbiosis. *Proc Natl Acad Sci U S A* **103**: 359-364.
- 789 Kim S, Soltis PS, Wall K, Soltis DE. 2006. Phylogeny and domain evolution in the
790 *APETALA2*-like gene family. *Mol Biol Evol* **23**: 107-120.
- 791 Kistner C, Winzer T, Pitzschke A, Mulder L, Sato S, Kaneko T, Tabata S,
792 Sandal N, Stougaard J, Webb KJ, Szczyglowski K, Parniske M. 2005. Seven

- 793 *Lotus japonicus* genes required for transcriptional reprogramming of the root during
794 fungal and bacterial symbiosis. *Plant Cell* **17**: 2217-2229.
- 795 **Kozomara A, Griffiths-Jones S. 2011.** miRBase: integrating microRNA annotation
796 and deep-sequencing data. *Nucleic Acids Res* **39**: D152-157.
- 797 **Krusell L, Krause K, Ott T, Desbrosses G, Kramer U, Sato S, Nakamura Y,**
798 **Tabata S, James EK, Sandal N, Stougaard J, Kawaguchi M, Miyamoto A,**
799 **Suganuma N, Udvardi MK. 2005.** The sulfate transporter SST1 is crucial for
800 symbiotic nitrogen fixation in *Lotus japonicus* root nodules. *Plant Cell* **17**: 1625-
801 1636.
- 802 **Krusell L, Madsen LH, Sato S, Aubert G, Genua A, Szczygłowski K, Duc G,**
803 **Kaneko T, Tabata S, de Bruijn F, Pajuelo E, Sandal N, Stougaard J. 2002.** Shoot
804 control of root development and nodulation is mediated by a receptor-like kinase.
805 *Nature* **420**: 422-426.
- 806 **Kulcheski FR, de Oliveira LF, Molina LG, Almerao MP, Rodrigues FA,**
807 **Marcolino J, Barbosa JF, Stolf-Moreira R, Nepomuceno AL, Marcelino-**
808 **Guimaraes FC, Abdelnoor RV, Nascimento LC, Carazzolle MF, Pereira GA,**
809 **Margis R. 2011.** Identification of novel soybean microRNAs involved in abiotic and
810 biotic stresses. *BMC Genomics* **12**: 307.
- 811 **Lauressergues D, Delaux PM, Formey D, Lelandais-Briere C, Fort S, Cottaz S,**
812 **Becard G, Niebel A, Roux C, Combier JP. 2012.** The microRNA miR171h
813 modulates arbuscular mycorrhizal colonization of *Medicago truncatula* by targeting
814 *NSP2*. *Plant J* **72**: 512-522.
- 815 **Lauter N, Kampani A, Carlson S, Goebel M, Moose SP. 2005.** microRNA172
816 down-regulates *glossy15* to promote vegetative phase change in maize. *Proc Natl*
817 *Acad Sci U S A* **102**: 9412-9417.
- 818 **Lelandais-Briere C, Naya L, Sallet E, Calenge F, Frugier F, Hartmann C, Gouzy**
819 **J, Crespi M. 2009.** Genome-wide *Medicago truncatula* small RNA analysis revealed
820 novel microRNAs and isoforms differentially regulated in roots and nodules. *Plant*
821 *Cell* **21**: 2780-2796.
- 822 **Lelandais-Briere C, Sorin C, Declerck M, Benslimane A, Crespi M, Hartmann**
823 **C. 2010.** Small RNA diversity in plants and its impact in development. *Curr*
824 *Genomics* **11**: 14-23.
- 825 **Levy J, Bres C, Geurts R, Chalhoub B, Kulikova O, Duc G, Journet EP, Ane**
826 **JM, Lauber E, Bisseling T, Denarie J, Rosenberg C, Debelle F. 2004.** A putative
827 Ca²⁺ and calmodulin-dependent protein kinase required for bacterial and fungal
828 symbioses. *Science* **303**: 1361-1364.

- 829 **Li H, Deng Y, Wu T, Subramanian S, Yu O. 2010.** Misexpression of miR482,
830 miR1512, and miR1515 increases soybean nodulation. *Plant Physiol* **153**: 1759-1770.
- 831 **Li X, Wang X, Zhang S, Liu D, Duan Y, Dong W. 2012.** Identification of soybean
832 microRNAs involved in soybean cyst nematode infection by deep sequencing. *PLoS*
833 *ONE* **7**: e39650.
- 834 **Limpens E, Bisseling T. 2014.** CYCLOPS: a new vision on rhizobium-induced
835 nodule organogenesis. *Cell Host Microbe* **15**: 127-129.
- 836 **Llave C, Xie Z, Kasschau KD, Carrington JC. 2002.** Cleavage of Scarecrow-like
837 mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* **297**: 2053-2056.
- 838 **Madsen EB, Antolin-Llovera M, Grossmann C, Ye J, Vieweg S, Broghammer A,**
839 **Krusell L, Radutoiu S, Jensen ON, Stougaard J, Parniske M. 2011.**
840 Autophosphorylation is essential for the in vivo function of the *Lotus japonicus* Nod
841 factor receptor 1 and receptor-mediated signalling in cooperation with Nod factor
842 receptor 5. *Plant J* **65**: 404-417.
- 843 **Madsen EB, Madsen LH, Radutoiu S, Olbryt M, Rakwalska M, Szczygłowski K,**
844 **Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J. 2003.** A receptor kinase gene
845 of the LysM type is involved in legume perception of rhizobial signals. *Nature* **425**:
846 637-640.
- 847 **Madsen LH, Tirichine L, Jurkiewicz A, Sullivan JT, Heckmann AB, Bek AS,**
848 **Ronson CW, James EK, Stougaard J. 2010.** The molecular network governing
849 nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nat*
850 *Commun* **1**: 10.
- 851 **Maekawa T, Maekawa-Yoshikawa M, Takeda N, Imaizumi-Anraku H, Murooka**
852 **Y, Hayashi M. 2009.** Gibberellin controls the nodulation signaling pathway in *Lotus*
853 *japonicus*. *Plant J* **58**: 183-194.
- 854 **Matsui A, Nguyen AH, Nakaminami K, Seki M. 2013.** Arabidopsis non-coding
855 RNA regulation in abiotic stress responses. *Int J Mol Sci* **14**: 22642-22654.
- 856 **Messinese E, Mun JH, Yeun LH, Jayaraman D, Rouge P, Barre A, Loughnon G,**
857 **Schornack S, Bono JJ, Cook DR, Ane JM. 2007.** A novel nuclear protein interacts
858 with the symbiotic DMI3 calcium- and calmodulin-dependent protein kinase of
859 *Medicago truncatula*. *Mol Plant Microbe Interact* **20**: 912-921.
- 860 **Middleton PH, Jakab J, Penmetsa RV, Starker CG, Doll J, Kalo P, Prabhu R,**
861 **Marsh JF, Mitra RM, Kereszt A, Dudas B, VandenBosch K, Long SR, Cook DR,**
862 **Kiss GB, Oldroyd GE. 2007.** An ERF transcription factor in *Medicago truncatula*
863 that is essential for Nod factor signal transduction. *Plant Cell* **19**: 1221-1234.

- 864 **Miwa H, Sun J, Oldroyd GE, Downie JA. 2006.** Analysis of Nod-factor-induced
865 calcium signaling in root hairs of symbiotically defective mutants of *Lotus japonicus*.
866 *Mol Plant Microbe Interact* **19**: 914-923.
- 867 **Molina LG, da Fonseca GC, de Moraes GL, de Oliveira LF, de Carvalho JB,**
868 **Kulcheski FR, Margis R. 2012.** Metatranscriptomic analysis of small RNAs present
869 in soybean deep sequencing libraries. *Genet Mol Biol* **35**: 292-303.
- 870 **Murakami Y, Miwa H, Imaizumi-Anraku H, Kouchi H, Downie JA, Kawaguchi**
871 **M, Kawasaki S. 2006.** Positional cloning identifies *Lotus japonicus* NSP2, a putative
872 transcription factor of the GRAS family, required for *NIN* and *ENOD40* gene
873 expression in nodule initiation. *DNA Res* **13**: 255-265.
- 874 **Murray JD, Karas BJ, Sato S, Tabata S, Amyot L, Szczylowski K. 2007.** A
875 cytokinin perception mutant colonized by *Rhizobium* in the absence of nodule
876 organogenesis. *Science* **315**: 101-104.
- 877 **Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O,**
878 **Jones JD. 2006.** A plant miRNA contributes to antibacterial resistance by repressing
879 auxin signaling. *Science* **312**: 436-439.
- 880 **Oldroyd GE. 2013.** Speak, friend, and enter: signalling systems that promote
881 beneficial symbiotic associations in plants. *Nat Rev Microbiol* **11**: 252-263.
- 882 **Oldroyd GE, Downie JA. 2006.** Nuclear calcium changes at the core of symbiosis
883 signalling. *Curr Opin Plant Biol* **9**: 351-357.
- 884 **Padmanabhan C, Zhang X, Jin H. 2009.** Host small RNAs are big contributors to
885 plant innate immunity. *Curr Opin Plant Biol* **12**: 465-472.
- 886 **Perry J, Brachmann A, Welham T, Binder A, Charpentier M, Groth M, Haage**
887 **K, Markmann K, Wang TL, Parniske M. 2009.** TILLING in *Lotus japonicus*
888 identified large allelic series for symbiosis genes and revealed a bias in functionally
889 defective EMS alleles towards glycine replacements. *Plant Physiol* **151**: 1281-1291.
- 890 **Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, Gronlund M, Sato**
891 **S, Nakamura Y, Tabata S, Sandal N, Stougaard J. 2003.** Plant recognition of
892 symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **425**: 585-592.
- 893 **Radutoiu S, Madsen LH, Madsen EB, Jurkiewicz A, Fukai E, Quistgaard EM,**
894 **Albrechtsen AS, James EK, Thirup S, Stougaard J. 2007.** LysM domains mediate
895 lipochitin-oligosaccharide recognition and *Nfr* genes extend the symbiotic host range.
896 *Embo J* **26**: 3923-3935.
- 897 **Ramachandran V, Chen X. 2008.** Small RNA metabolism in Arabidopsis. *Trends*
898 *Plant Sci* **13**: 368-374.

- 899 **Rodporthong P, Sullivan JT, Songsrirote K, Sumpton D, Cheung KW, Thomas-**
 900 **Oates J, Radutoiu S, Stougaard J, Ronson CW. 2009.** Nodulation gene mutants of
 901 *Mesorhizobium loti* R7A-*nodZ* and *noll* mutants have host-specific phenotypes on
 902 *Lotus spp.* *Mol Plant Microbe Interact* **22**: 1546-1554.
- 903 **Ruiz-Ferrer V, Voinnet O. 2009.** Roles of plant small RNAs in biotic stress
 904 responses. *Annu Rev Plant Biol* **60**: 485-510.
- 905 **Saito K, Yoshikawa M, Yano K, Miwa H, Uchida H, Asamizu E, Sato S, Tabata**
 906 **S, Imaizumi-Anraku H, Umehara Y, Kouchi H, Murooka Y, Szczygłowski K,**
 907 **Downie JA, Parniske M, Hayashi M, Kawaguchi M. 2007.** NUCLEOPORIN85 is
 908 required for calcium spiking, fungal and bacterial symbioses, and seed production in
 909 *Lotus japonicus*. *Plant Cell* **19**: 610-624.
- 910 **Sandal N, Petersen TR, Murray J, Umehara Y, Karas B, Yano K, Kumagai H,**
 911 **Yoshikawa M, Saito K, Hayashi M, Murakami Y, Wang X, Hakoyama T,**
 912 **Imaizumi-Anraku H, Sato S, Kato T, Chen W, Hossain MS, Shibata S, Wang**
 913 **TL, Yokota K, Larsen K, Kanamori N, Madsen E, Radutoiu S, Madsen LH,**
 914 **Radu TG, Krusell L, Ooki Y, Banba M, Betti M, Rispail N, Skot L, Tuck E,**
 915 **Perry J, Yoshida S, Vickers K, Pike J, Mulder L, Charpentier M, Muller J,**
 916 **Ohtomo R, Kojima T, Ando S, Marquez AJ, Gresshoff PM, Harada K, Webb J,**
 917 **Hata S, Suganuma N, Kouchi H, Kawasaki S, Tabata S, Hayashi M, Parniske M,**
 918 **Szczygłowski K, Kawaguchi M, Stougaard J. 2006.** Genetics of symbiosis in *Lotus*
 919 *japonicus*: recombinant inbred lines, comparative genetic maps, and map position of
 920 35 symbiotic loci. *Mol Plant Microbe Interact* **19**: 80-91.
- 921 **Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S,**
 922 **Watanabe A, Ono A, Kawashima K, Fujishiro T, Katoh M, Kohara M, Kishida**
 923 **Y, Minami C, Nakayama S, Nakazaki N, Shimizu Y, Shinpo S, Takahashi C,**
 924 **Wada T, Yamada M, Ohmido N, Hayashi M, Fukui K, Baba T, Nakamichi T,**
 925 **Mori H, Tabata S. 2008.** Genome structure of the legume *Lotus japonicus*. *DNA Res*
 926 **15**: 227-239.
- 927 **Schauser L, Handberg K, Sandal N, Stiller J, Thykjaer T, Pajuelo E, Nielsen A,**
 928 **Stougaard J. 1998.** Symbiotic mutants deficient in nodule establishment identified
 929 after T-DNA transformation of *Lotus japonicus*. *Mol Gen Genet* **259**: 414-423.
- 930 **Schauser L, Roussis A, Stiller J, Stougaard J. 1999.** A plant regulator controlling
 931 development of symbiotic root nodules. *Nature* **402**: 191-195.
- 932 **Shigyo M, Hasebe M, Ito M. 2006.** Molecular evolution of the AP2 subfamily. *Gene*
 933 **366**: 256-265.
- 934 **Shivaprasad PV, Chen HM, Patel K, Bond DM, Santos BA, Baulcombe DC.**
 935 **2012.** A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats
 936 and other mRNAs. *Plant Cell* **24**: 859-874.

- 937 **Singh S, Katzer K, Lambert J, Cerri M, Parniske M. 2014.** CYCLOPS, a DNA-
938 binding transcriptional activator, orchestrates symbiotic root nodule development.
939 *Cell Host Microbe* **15**: 139-152.
- 940 **Smit P, Raedts J, Portyanko V, Debelle F, Gough C, Bisseling T, Geurts R. 2005.**
941 NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced
942 transcription. *Science* **308**: 1789-1791.
- 943 **Soyano T, Kouchi H, Hirota A, Hayashi M. 2013.** Nodule inception directly targets
944 NF-Y subunit genes to regulate essential processes of root nodule development in
945 *Lotus japonicus*. *PLoS Genet* **9**: e1003352.
- 946 **Stracke S, Kistner C, Yoshida S, Mulder L, Sato S, Kaneko T, Tabata S, Sandal**
947 **N, Stougaard J, Szczyglowski K, Parniske M. 2002.** A plant receptor-like kinase
948 required for both bacterial and fungal symbiosis. *Nature* **417**: 959-962.
- 949 **Subramanian S, Fu Y, Sunkar R, Barbazuk WB, Zhu JK, Yu O. 2008.** Novel and
950 nodulation-regulated microRNAs in soybean roots. *BMC Genomics* **9**: 160.
- 951 **Suganuma N, Nakamura Y, Yamamoto M, Ohta T, Koiwa H, Akao S,**
952 **Kawaguchi M. 2003.** The *Lotus japonicus Sen1* gene controls rhizobial
953 differentiation into nitrogen-fixing bacteroids in nodules. *Mol Genet Genomics* **269**:
954 312-320.
- 955 **Szitty G, Moxon S, Santos DM, Jing R, Fevereiro MP, Moulton V, Dalmay T.**
956 **2008.** High-throughput sequencing of *Medicago truncatula* short RNAs identifies
957 eight new miRNA families. *BMC Genomics* **9**: 593.
- 958 **Tirichine L, Imaizumi-Anraku H, Yoshida S, Murakami Y, Madsen LH, Miwa**
959 **H, Nakagawa T, Sandal N, Albrechtsen AS, Kawaguchi M, Downie A, Sato S,**
960 **Tabata S, Kouchi H, Parniske M, Kawasaki S, Stougaard J. 2006a.** Deregulation
961 of a Ca²⁺/calmodulin-dependent kinase leads to spontaneous nodule development.
962 *Nature* **441**: 1153-1156.
- 963 **Tirichine L, James EK, Sandal N, Stougaard J. 2006b.** Spontaneous root-nodule
964 formation in the model legume *Lotus japonicus*: a novel class of mutants nodulates in
965 the absence of rhizobia. *Mol Plant Microbe Interact* **19**: 373-382.
- 966 **Tirichine L, Sandal N, Madsen LH, Radutoiu S, Albrechtsen AS, Sato S, Asamizu**
967 **E, Tabata S, Stougaard J. 2007.** A gain-of-function mutation in a cytokinin receptor
968 triggers spontaneous root nodule organogenesis. *Science* **315**: 104-107.
- 969 **Turner M, Nizampatnam NR, Baron M, Coppin S, Damodaran S, Adhikari S,**
970 **Arunachalam SP, Yu O, Subramanian S. 2013.** Ectopic expression of miR160
971 results in auxin hypersensitivity, cytokinin hyposensitivity, and inhibition of
972 symbiotic nodule development in soybean. *Plant Physiol* **162**: 2042-2055.

- 973 **Valdes-Lopez O, Yang SS, Aparicio-Fabre R, Graham PH, Reyes JL, Vance CP,**
 974 **Hernandez G. 2010.** MicroRNA expression profile in common bean (*Phaseolus*
 975 *vulgaris*) under nutrient deficiency stresses and manganese toxicity. *New Phytol* **187**:
 976 805-818.
- 977 **Vernie T, Moreau S, de Billy F, Plet J, Combier JP, Rogers C, Oldroyd G,**
 978 **Frugier F, Niebel A, Gamas P. 2008.** EFD is an ERF transcription factor involved in
 979 the control of nodule number and differentiation in *Medicago truncatula*. *Plant Cell*
 980 **20**: 2696-2713.
- 981 **Wang T, Chen L, Zhao M, Tian Q, Zhang WH. 2011.** Identification of drought-
 982 responsive microRNAs in *Medicago truncatula* by genome-wide high-throughput
 983 sequencing. *BMC Genomics* **12**: 367.
- 984 **Wang Y, Li P, Cao X, Wang X, Zhang A, Li X. 2009.** Identification and expression
 985 analysis of miRNAs from nitrogen-fixing soybean nodules. *Biochem Biophys Res*
 986 *Commun* **378**: 799-803.
- 987 **Wang Y, Wang L, Zou Y, Chen L, Cai Z, Zhang S, Zhao F, Tian Y, Jiang Q,**
 988 **Ferguson BJ, Gresshoff PM, Li X. 2014a.** Soybean miR172c Targets the Repressive
 989 AP2 Transcription Factor NNC1 to Activate ENOD40 Expression and Regulate
 990 Nodule Initiation. *Plant Cell* **26**: 4782-4801.
- 991 **Wang Y, Wang Z, Amyot L, Tian L, Xu Z, Gruber MY, Hannoufa A. 2014b.**
 992 Ectopic expression of miR156 represses nodulation and causes morphological and
 993 developmental changes in *Lotus japonicus*. *Mol Genet Genomics* **290**: 471-484.
- 994 **Wopereis J, Pajuelo E, Dazzo FB, Jiang Q, Gresshoff PM, De Bruijn FJ,**
 995 **Stougaard J, Szczyglowski K. 2000.** Short root mutant of *Lotus japonicus* with a
 996 dramatically altered symbiotic phenotype. *Plant J* **23**: 97-114.
- 997 **Wu G, Park MY, Conway SR, Wang JW, Weigel D, Poethig RS. 2009.** The
 998 sequential action of miR156 and miR172 regulates developmental timing in
 999 *Arabidopsis*. *Cell* **138**: 750-759.
- 1000 **Yamashino T, Yamawaki S, Hagui E, Ishida K, Ueoka-Nakanishi H, Nakamichi**
 1001 **N, Mizuno T. 2013.** Clock-controlled and FLOWERING LOCUS T (FT)-dependent
 1002 photoperiodic pathway in *Lotus japonicus* II: characterization of a microRNA
 1003 implicated in the control of flowering time. *Biosci Biotechnol Biochem* **77**: 1179-
 1004 1185.
- 1005 **Yan Z, Hossain MS, Wang J, Valdes-Lopez O, Liang Y, Libault M, Qiu L,**
 1006 **Stacey G. 2013.** miR172 regulates soybean nodulation. *Mol Plant Microbe Interact*
 1007 **26**: 1371-1377.
- 1008 **Yang X, Li L. 2011.** miRDeep-P: a computational tool for analyzing the microRNA
 1009 transcriptome in plants. *Bioinformatics* **27**: 2614-2615.

1010 **Yano K, Yoshida S, Müller J, Singh S, Banba M, Vickers K, Markmann K,**
1011 **White C, Schuller B, Sato S, Asamizu E, Tabata S, Murooka Y, Perry J, Wang**
1012 **T, Kawaguchi M, Imaizumi-Anraku H, Hayashi M, Parniske M. 2008a.**
1013 **CYCLOPS, a mediator of symbiotic intracellular accomodation. *Proc Natl Acad Sci***
1014 ***U S A* **105**: 20540-20545.**

1015 **Yano K, Yoshida S, Muller J, Singh S, Banba M, Vickers K, Markmann K,**
1016 **White C, Schuller B, Sato S, Asamizu E, Tabata S, Murooka Y, Perry J, Wang**
1017 **TL, Kawaguchi M, Imaizumi-Anraku H, Hayashi M, Parniske M. 2008b.**
1018 **CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proc Natl Acad Sci***
1019 ***U S A* **105**: 20540-20545.**

1020 **Yu S, Galvao VC, Zhang YC, Horrer D, Zhang TQ, Hao YH, Feng YQ, Wang S,**
1021 **Schmid M, Wang JW. 2012. Gibberellin regulates the Arabidopsis floral transition**
1022 **through miR156-targeted SQUAMOSA promoter binding-like transcription factors.**
1023 ***Plant Cell* **24**: 3320-3332.**

1024 **Zabala G, Campos E, Varala KK, Bloomfield S, Jones SI, Win H, Tuteja JH,**
1025 **Calla B, Clough SJ, Hudson M, Vodkin LO. 2012. Divergent patterns of**
1026 **endogenous small RNA populations from seed and vegetative tissues of *Glycine max*.**
1027 ***BMC plant biology* **12**: 177.**

1028 **Zhai J, Jeong DH, De Paoli E, Park S, Rosen BD, Li Y, Gonzalez AJ, Yan Z,**
1029 **Kitto SL, Grusak MA, Jackson SA, Stacey G, Cook DR, Green PJ, Sherrier DJ,**
1030 **Meyers BC. 2011. MicroRNAs as master regulators of the plant NB-LRR defense**
1031 **gene family via the production of phased, trans-acting siRNAs. *Genes Dev* **25**: 2540-**
1032 **2553.**

1033 **Zhu QH, Helliwell CA. 2011. Regulation of flowering time and floral patterning by**
1034 **miR172. *J Exp Bot* **62**: 487-495.**

1035 **Zhu QH, Upadhyaya NM, Gubler F, Helliwell CA. 2009. Over-expression of**
1036 **miR172 causes loss of spikelet determinacy and floral organ abnormalities in rice**
1037 **(*Oryza sativa*). *BMC plant biology* **9**: 149.**

1038

1039

1040 Supporting Information

1041 **Fig. S1.** Transgenic *L. japonicus* roots overexpressing *MIR172a*, a Short Tandem
1042 Target Mimic targeting miR172a (172STTM), *RAP2-7-like1* or a miR172a resistant
1043 version of *RAP2-7-like1*, form nodules similar to control roots in morphology and
1044 number.

1045 **Fig. S2.** *L. japonicus rap2-7-like1* mutant plants form nodules similar to control
1046 plants in morphology and number.

1047 **Fig. S3.** Model of the genetic dependence of miR172b induction in *L. japonicus* roots
1048 during early (3 days) infection with *M. loti*.

1049 **Table S1.** Oligonucleotide sequences used (a) in reverse transcription, (b) in
1050 quantitative PCR and (c) for cloning.

1051 **Table S2** Sequence length distribution of redundant and unique dataset.

1052 **Table S3a** Predicted novel miRNAs in our dataset, associated expression data and
1053 mapping results to *L. japonicus* genome v2.5.

1054 **Table S3b** Genomic loci generating predicted novel miRNA candidates, associated
1055 precursor sequence and secondary structure, and miRNA* sequences with associated
1056 expression information.

1057 **Table S3c** miRDeep-P (Yang & Li, 2011) predicted miRNAs in our dataset that map
1058 to (three-mismatch-criterium, miRBase v19) or are contained in the miRBase miRNA
1059 registry along with associated expression and mapping data.

1060 **Table S4** Novel and conserved miRNA candidates that show a more than 10x
1061 difference in abundance between mock and inoculated libraries, and cloned miR172
1062 isoforms.

1063 **Table S5** Sequences not predicted as miRNAs by mirDeep-P but mapping to
1064 conserved plant miRNAs (miRBase v19).

1065 **Table S6** Target prediction and validation results for miR172a.

1066 **Table S7** Affymetrix gene chip expression data for the *MIR172a* precursor gene and
1067 predicted miR172a targets.

1068 **Table S8** Several promoter-transgene combinations were examined in *Lotus*
1069 *japonicus* transgenic roots against an empty vector control, but none showed an
1070 apparent phenotypical effect.

1071 **Methods S1** Supporting Methods.

1072 Notes S1 Misexpression of *MIR172a* or its targets suggests a fine-tuning role.

1073

1074 **Tables**

Table 1 *Lotus japonicus* tissue samples used to generate
sRNA libraries.

sample #	replicate #	treatment
1	1	mock (3 dpi)
2	1	<i>M. loti</i> (3 dpi)
3	2	mock (3 dpi)
4	2	<i>M. loti</i> (3 dpi)

1075 sRNA, small RNA; *M. loti*, *Mesorhizobium loti*; dpi, days post inoculation.

1076

1077

Table 2 Unique sequences (19-24nt) responsive to *Lotus japonicus* inoculation with *Mesorhizobium loti*.

sample #	biological replicate	10-fold down	100-fold down	10-fold up	100-fold up
1 vs 2	1	553	44	1372	113
3 vs 4	2	734	68	1414	260
1 vs 2 AND 3 vs 4	1 and 2	456	14	1034	62

1078 Calculations are based on normalized sequence counts. Only sequences represented
 1079 with a minimum total of 100 normalized counts are considered. vs, *versus*.

1080

1081

1082 **Figure Legends**

1083 **Fig. 1.** Flow scheme of *L. japonicus* sRNA sequence data analysis. Numbers indicate
 1084 output details of the combined four sRNA libraries. Analyses were done using the
 1085 shortran pipeline (Gupta *et al.*, 2012).

1086 **Fig. 2.** Sequence length distribution of redundant and unique *L. japonicus* sRNA
 1087 datasets. Error bars indicate Standard Deviation between the libraries.

1088 **Fig. 3.** Sequence expression distribution. Sequences are plotted to reflect the log of
 1089 normalized counts in two of the libraries each. (a) Two sequences with a strong
 1090 abundance increase in infected roots are encircled. Arrowheads point out miR172a.
 1091 Left panel, replicate 1; right panel, replicate 2. (b) Global expression comparison
 1092 between replicates. Left panel, mock and right panel, *M. loti* treated *L. japonicus*
 1093 roots.

1094 **Fig. 4.** Abundance profiles of miRNA172a and miR171c in *L. japonicus* roots and
 1095 nodules. (Bancos *et al.*), Left panels, normalized sequence counts in Illumina
 1096 libraries. Numbers below graphs indicate the log(inoculated/uninoculated) of
 1097 normalized counts of two biological replicates. Right panels, qRT-PCR targeting the
 1098 mature sRNAs using the same RNAs used for sequencing. (c-d) qRT-PCR targeting
 1099 miRNA172a within (c) hours or (d) days post *M. loti* inoculation. Expression values
 1100 are relative to mock treated tissues of the respective type and time point. Error bars
 1101 show Standard Error of the Mean of three biological replicates. qRT-PCRs may co-
 1102 target other miR172 and miR171 family members, respectively. dpi, days post
 1103 inoculation; hpi, hours post inoculation. qRT-PCR, reverse transcription followed by
 1104 quantitative PCR.

1105 **Fig. 5.** Infection induced abundance increase of miR172a and *NIN* in *L. japonicus*
 1106 symbiotic mutants. (a-c) qRT-PCR targeting (a, b) the mature miRNA172a sequence
 1107 or (c) *NIN* mRNA. Expression values are relative to mock treated roots of each
 1108 respective genotype. Error bars show Standard Error of the Mean of three biological
 1109 replicates. d, days. qRT-PCR, reverse transcription followed by quantitative PCR.

1110 **Fig. 6.** miRNA172a induction in *L. japonicus* roots upon *M. loti* infection depends on
 1111 both bacteria and Nod factors. (a) Gifu wt, *nfr1* and *nfr5* mutant roots treated with

either *M. loti* strain R7A, *nodC* *M. loti* (R7A) secreting no Nod factor or purified Nod factor (R7A). (b) *NIN* and *NF-YAI* responses serve as controls. *NF-YAI* was near the lower detection limit in wild type mock roots, and not detected (n.d.) in *nfr1* and *nfr5* roots. (Bancos *et al.*) qRT-PCR targeting (a) miRNA172a, (b, left panels) *NIN* mRNA or (b, right panel) *NF-YAI* mRNA. Values are relative to mock treated tissues of the respective genotype. Error bars show Standard Error of the Mean of three biological replicates. d, days. qRT-PCR, reverse transcription followed by quantitative PCR.

Fig. 7. miR172a is highly expressed in *L. japonicus* nodules compared to roots, regardless of nitrogen fixation activity (*sst1* and *sen1*), and of bacterial presence (*snf1*). qRT-PCR expression values are relative to mock treated roots of respective genotypes. Right and central panels stem from the same experiment. Error bars show Standard Error of the Mean of three biological replicates. qRT-PCR, reverse transcription followed by quantitative PCR.

Fig. 8. *pMIR172a* promoter activity coincides with and precedes bacterial infection in *L. japonicus* roots and nodules. Roots/nodules harvested at (a-b) 1 day post inoculation (dpi), (c-d) 2 dpi, (e-f) 3 dpi, (g-h) 4 dpi, (i) and (k-m) 3 weeks post inoculation (wpi), (j) 7 dpi, (o) 4 dpi, (n, p) 9 dpi and (q) 2 wpi. (a-m) GUS stained roots and nodules expressing *pMIR172a:GUS*. (n-q) Roots and nodules expressing *pMIR172a:tYFPnls*. Green fluorescence and white arrows indicate *pMIR172a:tYFPnls* activity. Red fluorescence, *M. loti* MAFF 303099 expressing *dsRED*. Black arrow indicates infection thread. Scale bars represent (a, c, e, g) 250 μ m; 50 μ m (b, d, f, h, n); (i) 1 mm; (j-m) 200 μ m and (o-q) 100 μ m. Tissue sections: 70 μ m.

Fig. 9. The predicted *L. japonicus* AP2-like transcription factor targets are endonucleolytically cleaved within the miR172a complementary site *in vivo*. Alignment of *LjRAP2-7-like1*, *LjRAP2-7-like2*, *LjRAP2-7-like3*, *LjAP2-like1* and *LjAP2-like2* with miR172a at complementary site. Arrowheads and numbers indicate detected sites and counts of non-capped 5' mRNA ends/total number of 5' ends detected within the miR172a complementary region. Stars indicate the dominant cleavage sites.

Fig. 10. miR172 targets *LjRAP2-7-like1*, *LjAP2-like1* and *LjAP2-like2* show complementary regulation to miR172a in *L. japonicus* nodules. (a-f) qRT-PCR

1144 targeting *LjRAP2-7-like1*, *LjAP2-like1* and *LjAP2-like2* during different stages of *M.*
1145 *loti* infection of *Lotus* (a-c) nodule- and (d-f) root tissues. Expression levels are
1146 relative to those in uninoculated roots. Error bars show Standard Error of the Mean of
1147 three biological replicates. dpi, days post inoculation; wpi, weeks post inoculation.
1148 qRT-PCR, reverse transcription followed by quantitative PCR.

144x93mm (300 x 300 DPI)

131x114mm (300 x 300 DPI)

177x407mm (300 x 300 DPI)

119x175mm (300 x 300 DPI)

77x74mm (300 x 300 DPI)

miR172a

48x28mm (300 x 300 DPI)

168x164mm (300 x 300 DPI)

128x189mm (300 x 300 DPI)

138x121mm (300 x 300 DPI)